

A New Approach to the Self-Dual Yang–Mills Equations

K. Takasaki*

Department of Mathematics, University of Tokyo, Hongo, Tokyo 113, Japan

Abstract. Inspired by Sato's new theory for soliton equations, we find a new approach to the self-dual Yang–Mills equations. We first establish a correspondence of solutions between the self-dual Yang–Mills equations and a new system of equations with infinitely many unknown functions. It then turns out that the latter equations can be easily solved by a simple explicit procedure. This leads to an explicit description of a very broad class of solutions to the self-dual Yang–Mills equations, and also to a construction of transformations acting on these solutions.

Introduction

Recently there has been significant progress due to Sato [1] in the theories of completely integrable systems. Further developments from slightly different aspects can be found in the papers of Date et al. [2], Mulase [3], Segal and Wilson [4], Ueno and Takasaki [5], etc. Sato's theory provides, in a unified framework, a new formulation of the "complete integrability" of various kinds of "soliton equations" (the KdV equation, the nonlinear Schrödinger equation, the Kadomtsev–Petviashvili equation, etc.), without using the conventional inverse scattering techniques including the Riemann–Hilbert problem. In his formulation all the information contained in a soliton equation is "coded" into the time evolution of $\infty \times \infty$ matrices which are regarded as points of an infinite dimensional Grassmann manifold, and in such a picture the structure of both solutions and transformations acting on them can be clearly described. In this paper we shall discuss a similar approach to the self-dual Yang–Mills equations.

In the four dimensional complex flat space \mathbb{C}^4 with coordinates $x = (y, z, \bar{y}, \bar{z})$, the (complexified) self-dual Yang–Mills equations with structure group $GL(r, \mathbb{C})$

* Address from April 1984: Department of Mathematics, Saitama University, 255 Shimo-Okubo, Urawa, Saitama 338, Japan

($r \geq 2$) are written in the form

$$\begin{aligned} [\partial_y + A_y, \partial_z + A_z] &= 0, [\partial_{\bar{y}} + A_{\bar{y}}, \partial_{\bar{z}} + A_{\bar{z}}] = 0, \\ [\partial_y + A_y, \partial_{\bar{y}} + A_{\bar{y}}] + [\partial_z + A_z, \partial_{\bar{z}} + A_{\bar{z}}] &= 0, \end{aligned} \quad (0.1)$$

where $\partial_u = \partial/\partial u$, $u = y, z, \bar{y}, \bar{z}$, and A_u , $u = y, z, \bar{y}, \bar{z}$, denote the $\mathcal{g}^\ell(r, \mathbb{C})$ -valued unknown functions depending on x . Contrary to the usual formulation, no reality conditions are imposed here. Then A_y and A_z can be eliminated by a suitable complex gauge transformation $A_u \rightarrow G^{-1}A_uG + G^{-1}\partial_uG$, $u = y, z, \bar{y}, \bar{z}$, $G = G(x)$, so that Eqs. (0.1) are reduced to

$$\partial_{\bar{y}}A_z - \partial_zA_{\bar{y}} + [A_{\bar{y}}, A_z] = 0, \partial_{\bar{y}}A_{\bar{y}} + \partial_zA_z = 0. \quad (0.2)$$

In what follows we shall mainly consider Eqs. (0.2), especially their formal power series solutions, $A_{\bar{y}}, A_z \in \mathcal{g}^\ell(r, \mathbb{C}[[x]])$, including local holomorphic solutions defined at $x = 0$.

It has been known that the self-dual Yang–Mills equations admit two different approaches—one from the twistor theory and the other from the inverse scattering method. The twistor approach, initiated by Ward [6], was combined with the algebraic geometry of vector bundles over the complex projective space \mathbb{P}^3 to determine the structure of all the instanton solutions (see Atiyah and Ward [7], Atiyah et al. [8], Atiyah [9] and the references therein). The inverse scattering approach, pointed out first by Belavin and Zakharov [10], was developed by using the Riemann–Hilbert problem to yield several results; the general construction of the Atiyah–Ward Ansatz by Corrigan et al. [11], the systematic derivation of the hidden symmetries by Ueno and Nakamura [12], etc. (see also Chau et al. [13], Dolan [14], Pohlmeyer [15], and Wu [16] for related topics), etc. Also it has been shown that these two approaches are closely related to each other (see Corrigan et al. [11] and Chau et al. [14]).

Our approach is based on a different principle (though closely related with the inverse scattering approach). We introduce, following the idea of [1], an $\infty \times \infty$ matrix $\xi = (\xi_{ij})_{i \in \mathbb{Z}, j < 0}$ consisting of $\mathcal{g}^\ell(r, \mathbb{C})$ -valued unknown functions ξ_{ij} (where \mathbb{Z} denotes the set of all integers), and consider the equations

$$\begin{aligned} -\partial_y \xi_{i+1,j} + \partial_z \xi_{ij} + \xi_{i-1} \partial_y \xi_{0j} &= 0, \partial_z \xi_{i+1,j} + \partial_{\bar{y}} \xi_{ij} - \xi_{i-1} \partial_z \xi_{0j} = 0, \\ \xi_{i+1,j} &= \xi_{i,j-1} + \xi_{i-1} \xi_{0j} \text{ for } i \in \mathbb{Z}, j < 0, \\ \xi_{ij} &= \delta_{ij} 1_r \text{ for } i, j < 0. \end{aligned} \quad (0.3)$$

Here 1_r denote the $r \times r$ unit matrix, and δ_{ij} Kronecker's delta. We first show, in the realm of both formal power series solutions and local holomorphic solutions, that the equations

$$A_{\bar{y}} = \partial_z \xi_{0-1}, A_z = -\partial_y \xi_{0-1} \quad (0.4)$$

define a correspondence of solutions (this is not a one-to-one correspondence) between Eqs. (0.2) and (0.3). To be more precise, we show that Eqs. (0.4) yield all the formal power series solutions to (0.2) from those to (0.3), and also all the local holomorphic solutions to (0.2) defined at $x = 0$ by imposing some analytical conditions on the solutions to (0.3). On the other hand, it turns out that we can

explicitly construct both these solutions to (0.3) and transformations acting on them by means of some simple manipulations of $\infty \times \infty$ matrices related with ξ . This construction is very similar to that of [1] for soliton equations, but there are also some decisive differences. Combining these facts, we can specify the structure of solutions to (0.2).

We shall develop our discussions as follows. In Sect. 1 we shall show that Eqs. (0.4) actually define a correspondence of solutions between Eqs. (0.2) and (0.3). We establish this fact by utilizing the linear system of Belavin and Zakharov [10], which we reinterpret as nonlinear differential equations connecting (0.2) and (0.3). Also the characteristic matrices originally introduced by Jimbo and Miwa [18] play an important role in these discussions. In Sect. 3 we shall explicitly solve (0.3). Our main theorems concerning the structure of solutions to (0.2) and (0.3) are stated here. (In the appendix a related topic is supplemented.) In Sect. 4 we shall discuss transformations acting on solutions to (0.3). (To be interesting enough, the expressions of these transformations and the associated infinitesimal transformations have forms similar to those of Hauser and Ernst [19] and Ueno and Nakamura [12].) Two classes of transformations, called type I and type II, are discussed in detail. The transformations of type I are shown to constitute a group, which transitively acts on the space of all the formal power series solutions. In a simplified case of transformations of type II there appears a hierarchy of differential equations which are similar to the AKNS systems [20] and consistently coupled with the self-dual Yang–Mills equations. In Sect. 4 we shall add a few comments on the relation of our method to Mulase’s method [3] and the Riemann–Hilbert problem.

List of Notations Used Throughout this Paper

\mathbb{C} and \mathbb{Z} denote respectively the sets of all complex numbers and all integers. $\mathbb{C}[[x]]$ and $\mathbb{C}[[y, z]]$ denote respectively the sets of all formal powers series of (y, z, \bar{y}, \bar{z}) and (y, z) (with coefficients in \mathbb{C}). $\mathcal{M}(r, R)$ ($R = \mathbb{C}, \mathbb{C}[[y, z]], \mathbb{C}[[x]]$, etc.) denotes the set of all $r \times r$ matrices with components in R . For a nonnegative

integer n and an element $f = \sum_{i,j,m=0}^{\infty} f_{ijkm} y^i z^j \bar{y}^k \bar{z}^m$ of $\mathbb{C}[[x]]$ or $\mathcal{M}(r, \mathbb{C}[[x]])$,

the notation $\text{ord}_x f \geq n$ means that $f_{ijkm} = 0$ for $i + j + k + m < n$. 1_r denotes the $r \times r$ unit matrix, and $\mathbb{1}$ the $\infty \times \infty$ unit matrix of various size. In this paper we shall freely utilize $\infty \times \infty$ matrices consisting of $r \times r$ blocks. Such a matrix is expressed in the form $(a_{ij})_{i \in I, j \in J}$, where I and J are some subsets of \mathbb{Z} , and a_{ij} (which we call the “ (ij) block” of this $\infty \times \infty$ matrix) is an $r \times r$ matrix. In our convention the blocks of these matrices are arrayed in such a way as indicated in the following example with $I = J = \mathbb{Z}$:

$$(a_{ij})_{i,j \in \mathbb{Z}} = \begin{pmatrix} \dots & \dots & \dots & \dots & \dots \\ \dots & a_{-1-1} & a_{-10} & a_{-11} & \dots \\ \dots & a_{0-1} & a_{00} & a_{01} & \dots \\ \dots & a_{1-1} & a_{10} & a_{11} & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

A denotes the $\infty \times \infty$ matrix $(\delta_{i+1,j} 1_r)_{i,j \in \mathbb{Z}}$, where δ_{ij} denotes Kronecher's delta. Finally, $\binom{a}{b}$ denotes the binomial coefficient, i.e. $\binom{a}{b} = a!/(b!(a-b)!)$.

1. Correspondence of Solutions between (0.2) and (0.3)

In this section we develop preliminary discussions about Eqs. (0.2) and related equations. The main purpose is to establish a correspondence of solutions between (0.2) and (0.3). The discussions developed below are divided into two steps.

The first step is to modify the observation of Belavin and Zakharov [10]. As pointed out in [10], Eqs. (0.2), which we may rewrite (using an indeterminate variable λ) in the following form

$$[-\lambda \partial_y + \partial_z + A_z, \lambda \partial_z + \partial_{\bar{y}} + A_{\bar{y}}] = 0,$$

are nothing but the integrability conditions of the linear system

$$(-\lambda \partial_y + \partial_z + A_z)W = 0, (\lambda \partial_z + \partial_{\bar{y}} + A_{\bar{y}})W = 0. \quad (1.1)$$

(To tell the truth, Belavin and Zakharov used a different gauge. The gauge adopted here is due to Chau et al. [13, 17] and Pohlmeyer [15].) In our case where formal power series solutions to (0.2) are considered, we require $W = W(x, \lambda)$ to be matrix-valued formal power series of the form

$$W = \sum_{j=0}^{\infty} W_j \lambda^{-j}, W_j \in \mathcal{gl}(r, \mathbb{C}[[x]]), W_0 = 1_r.$$

In terms of W_j Eqs. (1.1) are written in the form

$$-\partial_y W_{j+1} + \partial_z W_j + A_z W_j = 0, \partial_z W_{j+1} + \partial_{\bar{y}} W_j + A_{\bar{y}} W_j = 0, j \geq 0, \quad (1.2)$$

and recursively (though not uniquely) we can solve Eqs. (1.2), whose integrability conditions in the sense of Frobenius are assured by (0.2). (Note that this procedure is nothing but the well known construction of conservation laws. See, for example, Sect. III of [17]). It is also easy to show, conversely, that Eqs. (1.2) imply Eqs. (0.2). Thus we have

Proposition 1. *For any formal power series solution to (0.2), there is a solution to (1.2) with $W_j \in \mathcal{gl}(r, \mathbb{C}[[x]])$ and $W_0 = 1_r$. Conversely, if Eqs. (1.2) have such a solution, then Eqs. (0.2) are satisfied.*

When we consider local holomorphic solutions to (0.2), the usual argument for the integration of (1.1) (see, for example, Sect. 2 of [11]) shows

Proposition 2. *For any local holomorphic solution to (0.2) defined in a neighborhood of $x=0$, there is a local solution $W = W(x, \lambda)$ to (1.1) which is defined in a neighborhood of $(x, \lambda) = (0, \infty)$ in $\mathbb{C}^4 \times \mathbb{P}^1$, holomorphic with respect to (x, λ) , and satisfies the condition $W(x, \infty) = 1_r$.*

The equations for $j=0$ in (1.2) shows that $A_{\bar{y}}$ and A_z are recovered by

$$A_{\bar{y}} = -\partial_z W_1, A_z = \partial_y W_1. \quad (1.3)$$

Substituting (1.3) into (1.2), we finally obtain the equations

$$-\partial_y W_{j+1} + \partial_z W_j + (\partial_y W_1)W_j = 0, \partial_z W_{j+1} + \partial_y W_j - (\partial_z W_1)W_j = 0, j \geq 0, \quad (1.4)$$

for the new dependent variables $W_j, j \geq 0$. Proposition 1 shows that Eqs. (1.3) define a correspondence of solutions between Eqs. (0.2) and (1.4), and that all the formal power series solutions to (0.2) are derived from those to (1.4) via this correspondence.

The second step is to introduce a one-to-one correspondence of solutions between (1.4) and (0.3).

Let us first consider, apart from the differential equations, a one-to-one correspondence between a matrix-valued formal power series $W = \sum_{j=0}^{\infty} W_j \lambda^{-j}$ (with $W_j \in \mathfrak{gl}(r, \mathbb{C}[[x]])$ and $W_0 = 1_r$) and an $\infty \times \infty$ matrix $\xi = (\xi_{ij})_{i \in \mathbb{Z}, j < 0}$ (with $\xi_{ij} \in \mathfrak{gl}(r, \mathbb{C}[[x]])$) with the conditions

$$\Lambda \xi = \xi C, \quad (1.5)$$

$$\xi_{ij} = \delta_{ij} 1_r \quad \text{for } i, j < 0, \quad (1.6)$$

where $\Lambda = (\delta_{i+1, j} 1_r)_{i, j \in \mathbb{Z}}$ and

$$C = \begin{pmatrix} (\delta_{i+1, j} 1_r)_{i < -1, j < 0} \\ (\xi_{0j})_{j < 0} \end{pmatrix}. \quad (1.7)$$

(Note that (1.5) and (1.6) are the same as the third and fourth equations in (0.3).) The correspondence $W \leftrightarrow \xi$ between W and ξ is defined by

$$\xi_{0j} = -W_{-j}, j < 0. \quad (1.8)$$

(This actually defines a one-to-one correspondence because ξ is completely determined by the blocks $\xi_{0j}, j < 0$, by virtue of (1.5) and (1.6).) The structure of ξ is specified as follows.

Proposition 3. *We have*

$$\xi = (W_{i-j}^*)_{i \in \mathbb{Z}, j < 0} (W_{i-j})_{i, j < 0}, \quad (1.9)$$

where we set $W_j = W_j^* = 0$ for $j < 0$, and $W_j^*, j \geq 0$, denote the coefficients of W^{-1} , i.e. $W^{-1} = \sum_{j=0}^{\infty} W_j^* \lambda^{-j}$.

Note that $W_j^*, j \geq 0$, can be recursively calculated by the formula

$$W_j^* = - \sum_{k=0}^{j-1} W_k^* W_{j-k} + \delta_{j0} 1_r. \quad (1.10)$$

Proof of Proposition 3. Let $\xi' = (\xi'_{ij})_{i \in \mathbb{Z}, j < 0}$ denote the right-hand side of (1.9). To show $\xi = \xi'$, we have only to verify that a) (1.5) and (1.6) are also satisfied by ξ' in place of ξ , and b) $\xi_{0j} = \xi'_{0j}$ for any $j < 0$. Statement (b) follows immediately from (1.10). Also from the definition of W_j^* or (1.10),

$$\begin{aligned} (W_{i-j}^*)_{i, j < 0} &= (W_{i-j})_{i, j < 0}^{-1}, \\ (W_{i-j}^*)_{i, j < 0} \Lambda_{(-)} (W_{i-j})_{i, j < 0} &= C, \end{aligned}$$

where $A_{(-)} = (\delta_{i+1,j} 1_{r})_{i,j < 0}$, so that

$$A\xi' = (W_{i-j}^*)_{i \in \mathbb{Z}, j < 0} A_{(-)} (W_{i-j})_{i,j < 0} = \xi' C.$$

Thus statement (a) follows. This proves the proposition.

It turns out that the blocks $\xi_{ij}, i \geq 0, j < 0$, coincide with the characteristic matrices of $W = W(x, \lambda)$ introduced by Jimbo and Miwa [18]. Recall that the characteristic matrices $G_{ij}, i \geq 1, j \geq 1$, are originally defined by the equation

$$W(x, \mu)^{-1} W(x, \lambda) - 1_r = (\lambda - \mu) \sum_{i,j=1}^{\infty} G_{ij} \mu^{-1} \lambda^{-j} \quad (1.11)$$

of generating functions with indeterminate variables λ and μ .

Proposition 4. *We have*

$$\xi_{ij} = G_{i+1, -j} \quad \text{for } i \geq 0, j < 0. \quad (1.12)$$

Proof. Equating the coefficients of monomials of λ and μ in (1.11), we have

$$W_i^* = G_{i1}, W_j = -G_{1j}, W_i^* W_j = G_{i,j+1} - G_{i+1,j}, \quad (1.13)$$

while from (1.9) we have

$$\xi_{i-1, -j} = \sum_{k=0}^{j-1} W_{i+j-k-1}^* W_k \quad \text{for } i, j \geq 1.$$

Hence

$$\begin{aligned} \xi_{i-1, -j} &= W_{i+j-1}^* + \sum_{k=1}^{j-1} (G_{i+j-k-1, k+1} - G_{i+j-k, k}) \\ &= W_{i+j-1}^* + G_{ij} - G_{i+j-1, 1} \\ &= G_{ij}. \end{aligned}$$

This proves the proposition.

Now we can rewrite (1.4) in terms of ξ .

Proposition 5. *Via the correspondence $W \leftrightarrow \xi$ Eqs. (1.4) are equivalent to*

$$(-A\partial_y + \partial_z)\xi + \xi A = 0, (A\partial_z + \partial_y)\xi + \xi B = 0, \quad (1.14)$$

where

$$A = \begin{pmatrix} 0 \\ (\partial_y \xi_{0j})_{j < 0} \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ (-\partial_z \xi_{0j})_{j < 0} \end{pmatrix}. \quad (1.15)$$

Note that Eqs. (1.14) are the same as the first and second equations in (0.3). Hence this proposition shows the equivalence of (1.14) and (0.3) via the correspondence $W \leftrightarrow \xi$.

Proof of Proposition 5. By virtue of (1.12) we can rewrite (1.14) into the equations

$$\begin{aligned} -\partial_y G_{i+1,j} + \partial_z G_{ij} + G_{i1} \partial_y G_{1j} &= 0, \\ \partial_z G_{i+1,j} + \partial_{\bar{y}} G_{ij} - G_{i1} \partial_z G_{1j} &= 0. \end{aligned}$$

Using generating functions of G_{ij} with indeterminate variables λ and μ , we can once more rewrite the above equations into

$$\begin{aligned} (-\mu\partial_y + \partial_z) \sum_{i,j=1}^{\infty} G_{ij}\mu^{-i}\lambda^{-j} + \left(1_r + \sum_{i=1}^{\infty} G_{i1}\mu^{-i}\right) \partial_y \sum_{j=1}^{\infty} G_{1j}\lambda^{-j} &= 0, \\ (\mu\partial_z + \partial_{\bar{y}}) \sum_{i,j=1}^{\infty} G_{ij}\mu^{-i}\lambda^{-i} - \left(1_r + \sum_{i=1}^{\infty} G_{i1}\mu^{-i}\right) \partial_z \sum_{j=1}^{\infty} G_{1j}\lambda^{-j} &= 0. \end{aligned}$$

Finally, multiplying them by $\lambda - \mu$ and using (1.11) and (1.13), we obtain the equations

$$\begin{aligned} (-\lambda\partial_y + \partial_z)W(x, \lambda) \cdot W(x, \lambda)^{-1} &= (-\mu\partial_y + \partial_z)W(x, \mu) \cdot W(x, \mu)^{-1}, \\ (\lambda\partial_z + \partial_{\bar{y}})W(x, \lambda) \cdot W(x, \lambda)^{-1} &= (\mu\partial_z + \partial_{\bar{y}})W(x, \mu) \cdot W(x, \mu)^{-1}, \end{aligned}$$

which give equivalent expressions of (1.14). Both sides of these equations are consequently independent of λ and μ . Hence, defining $A_{\bar{y}}$ and A_z by them respectively, we conclude that (1.4) and (1.14) are equivalent to each other. This proves the proposition.

Thus we have shown that Eqs. (1.3) and (1.8) define a correspondence of solutions between (0.2) and (0.3), and that all the formal power series solutions to (0.2) are derived from those to (0.3) via this correspondence. Further, when we consider local holomorphic solutions to (0.2), we may suppose that the dependent variables ξ_{ij} are matrix-valued holomorphic functions with some quantitative conditions implied by Proposition 2. In this way, although this is not a one-to-one correspondence (because the solution to (1.2) is not uniquely determined by $A_{\bar{y}}$ and A_z), our original problem concerning the description of solutions to (0.2) is now converted into the problem of solving (0.3).

Let us conclude this section with a few important remarks on the equations discussed in this section.

Remarks. i) The initial value problem with respect to the two dimensional plane $\bar{y} = \bar{z} = 0$ provides a convenient framework for considering the structure of solutions to (0.3). In fact, it is not hard to show in the realm of formal power series solutions that any solution to (0.3) is uniquely determined by the initial value $\xi^{(0)} = \xi|_{\bar{y}=\bar{z}=0}$, which is consequently required to satisfy (1.5) and (1.6) in place of ξ . Thus the next problem is to describe the evolution $\xi^{(0)} \rightarrow \xi$. This problem will be solved in Sect. 2.

ii) A similar aspect is also true of (1.4). Namely, any formal power series solution to (1.4) is uniquely determined by the initial value $W^{(0)} = W|_{\bar{y}=\bar{z}=0}$. In Sect. 4 we shall discuss a method which directly characterizes the solution to this initial value problem. (Needless to say, this initial value problem and that mentioned in the above remark (i) are equivalent to each other via the correspondence $W \leftrightarrow \xi$ and the correspondence $W^{(0)} \leftrightarrow \xi^{(0)}$ induced to the initial values. Also note here that the formulas (1.8)–(1.13) are also valid for the initial values $\xi^{(0)}$, $W_j^{(0)} = W|_{\bar{y}=\bar{z}=0}$, $W_j^{(0)*} = W_j^*|_{\bar{y}=\bar{z}=0}$ and $G_{ij}^{(0)} = G_{ij}|_{\bar{y}=\bar{z}=0}$. We shall later utilize these formulas.)

iii) The coupled system of (1.5) and (1.14) have the following significant

properties (all the discussions in Sects. 2 and 3 crucially depend on them). First, if we forget conditions (1.6), this system admits the transformations of solutions of the form

$$\begin{aligned} A &\rightarrow H^{-1}AH + H^{-1}C\partial_y H - H^{-1}\partial_z H, & B &\rightarrow H^{-1}BH - H^{-1}C\partial_z H - H^{-1}\partial_y H, \\ C &\rightarrow H^{-1}CH, & \xi &\rightarrow \xi H, \end{aligned}$$

as far as H is an invertible $\infty \times \infty$ matrix depending on x and all the transformed quantities can be defined as $\infty \times \infty$ matrices in a suitable meaning. We can check this fact by simple calculations. Note, however, that formulas (1.7) and (1.15) are not preserved under these transformations; in the case where conditions (1.6) are not satisfied, (1.7) and (1.15) are not valid anymore. We can recover (1.6) by replacing ξ by $\xi(\xi_{ij})_{i,j < 0}^{-1}$ (this is one of the above transformation) on the assumption that $(\xi_{ij})_{i,j < 0}$ is invertible in a suitable meaning. Then (1.7) and (1.15) turn out to be valid again. (To see this, we have only to equate the ‘‘upper half’’ parts of the terms in (1.5) and (1.14). For example, we have

$$0 = (-\Lambda\partial_y + \partial_z)\xi + \xi A = \left(\begin{pmatrix} 0 \\ \dots \\ (-\partial_y \xi_{0j})_{j < 0} \\ \dots \end{pmatrix} + A \right),$$

which shows the first formula in (1.15).)

2. Construction of Solutions

In this section we discuss the structure of solutions to (0.3) by explicitly solving the initial value problem mentioned in Remark (i) at the end of Section 1.

2.1.

We now state our main theorems concerning the structure of solutions to (0.3) in the framework of the above initial value problem

We first consider the formal power series solutions. Let $\xi^{(0)} = (\xi_{ij}^{(0)})_{i \in \mathbb{Z}, j < 0}$ (with $\xi_{ij}^{(0)} \in \mathcal{gl}(r, \mathbb{C}[[y, z]])$) be an $\infty \times \infty$ matrix satisfying (1.5) and (1.6) in place of ξ . We can describe the evolution $\xi^{(0)} \rightarrow \xi$ as the following astonishingly simple manipulations of $\infty \times \infty$ matrices.

Theorem 6. Let $\tilde{\xi} = (\xi_{ij})_{i \in \mathbb{Z}, j < 0}$ and $\tilde{\xi}_{(\pm)}$ denote the $\infty \times \infty$ matrices (consisting of $r \times r$ blocks $\tilde{\xi}_{ij} \in \mathcal{gl}(r, \mathbb{C}[[x]])$) defined by

$$\begin{aligned} \tilde{\xi} &= \exp(\bar{z}\Lambda\partial_y - \bar{y}\Lambda\partial_z)\xi^{(0)} = \sum_{k=0}^{\infty} (\bar{z}\Lambda\partial_y - \bar{y}\Lambda\partial_z)^k \xi^{(0)}/k!, \\ \tilde{\xi}_{(+)} &= (\tilde{\xi}_{ij})_{i \geq 0, j < 0}, & \tilde{\xi}_{(-)} &= (\tilde{\xi}_{ij})_{i, j < 0}. \end{aligned} \tag{2.1}$$

Then the inverse $\tilde{\xi}_{(-)}^{-1}$ and the product $\tilde{\xi}_{(+)}\tilde{\xi}_{(-)}^{-1}$ can be defined as $\infty \times \infty$ matrices consisting of $r \times r$ blocks $\in \mathcal{gl}(r, \mathbb{C}[[x]])$, and the matrix

$$\xi = \tilde{\xi}_{(-)}^{-1} \tilde{\xi}_{(+)} = \begin{pmatrix} (\delta_{ij}1_r)_{i, j < 0} \\ \tilde{\xi}_{(+)} \\ \tilde{\xi}_{(-)}^{-1} \end{pmatrix} \tag{2.2}$$

satisfies (0.3) and the initial condition $\xi|_{y=z=0} = \xi^{(0)}$.

In the appendix we shall discuss more explicit expressions of the components of $\tilde{\xi}_{\xi_{(+)}\xi_{(-)}^{-1}}$.

We next consider the case where local holomorphic solutions to (0.2) defined in some neighborhoods of $x = 0$ are concerned. Let $\xi^{(0)}$ and ξ be the same as above, and recall the correspondences $W^{(0)} \leftrightarrow \xi^{(0)}$ and $W \leftrightarrow \xi$ mentioned in Remark (ii) at the end of Sect. 1. Of course we know from Proposition 5 that W gives solution to (1.1). Further, we have

Theorem 7. *Suppose that $W^{(0)} = \sum_{j=0}^{\infty} W_j^{(0)} \lambda^{-j}$ is holomorphic with respect to (y, z, λ) in a neighborhood of $(y, z, \lambda) = (0, 0, \infty)$ in $\mathbb{C}^2 \times \mathbb{P}^1$. Then $W = \sum_{j=0}^{\infty} W_j \lambda^{-j}$ obtained via the evolution $\xi^{(0)} \rightarrow \xi$ is also holomorphic with respect to (x, λ) in a neighborhood of $(x, \lambda) = (0, \infty)$ in $\mathbb{C}^4 \times \mathbb{P}^1$.*

As an immediate corollary of these theorems and the results of Section 1, we get (via (1.3) and (1.8)) the explicit description of both all the formal power series solutions and all the local holomorphic solutions to (0.2).

Our construction of solutions by means of (2.1) and (2.2) is very similar to that of Sato [1] for soliton equations. The main difference is that in our case there appears an operator $\exp(zA\partial_y - \bar{y}A\partial_z)$, whose counterpart in the discussions of [1] is merely an exponential function of an $\infty \times \infty$ -matrix-valued linear function.

2.2

We here prove Theorem 6.

First, let us consider how to justify the inverse $\tilde{\xi}_{(-)}^{-1}$ and the product $\tilde{\xi}_{(+)}\tilde{\xi}_{(-)}^{-1}$. According to the decomposition

$$\xi^{(0)} = \begin{pmatrix} \mathbb{1} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ (\xi_{ij}^{(0)})_{i \geq 0, j < 0} \end{pmatrix}, \quad \mathbb{1} = (\delta_{ij} \mathbb{1}_{r})_{i, j < 0},$$

$\tilde{\xi}$ is written in the form

$$\begin{aligned} \tilde{\xi} &= \begin{pmatrix} \mathbb{1} \\ 0 \end{pmatrix} + \tilde{\xi}, \\ \tilde{\xi} &= (\tilde{\xi}_{ij})_{i \in \mathbb{Z}, j < 0} = \exp(\bar{z}A\partial_y - \bar{y}A\partial_z) \begin{pmatrix} 0 \\ (\xi_{ij}^{(0)})_{i \geq 0, j < 0} \end{pmatrix} \end{aligned} \tag{2.3}$$

It is easy to show that $\tilde{\xi}_{ij} \in \mathcal{g}^{\ell} [r, \mathbb{C}[[x]]]$ and

$$\text{ord}_x \tilde{\xi}_{ij} \geq -i \quad \text{for } i, j < 0. \tag{2.4}$$

Hence we can define the k^{th} power $\tilde{\xi}_{(-)}^k$ of $\tilde{\xi}_{(-)} = (\tilde{\xi}_{ij})_{i, j < 0}$ as an $\infty \times \infty$ matrix consisting of $r \times r$ blocks $\in \mathcal{g}^{\ell} (r, \mathbb{C}[[x]])$ such that ord_x [the (ij) block of $\tilde{\xi}_{(-)}^k] \geq k - 1 - i$ for $i, j < 0$. By virtue of this fact we can justify the Neumann series $\mathbb{1} + \sum_{k=1}^{\infty} (-\tilde{\xi}_{(-)})^k$ in the same meaning, and we have

$$\tilde{\xi}_{(-)}^{-1} = \mathbb{1} + \sum_{k=1}^{\infty} (-\tilde{\xi}_{(-)})^k, \quad (2.5)$$

$$\text{ord}_x [\text{the } (ij)\text{ block of } \tilde{\xi}_{(-)}^{-1} - \mathbb{1}] \geq -i \text{ for } i, j < 0. \quad (2.6)$$

Further, (2.6) shows that we can also define the product $\tilde{\xi}_{(+)}\tilde{\xi}_{(-)}^{-1}$ in the same meaning. (In these discussions we must deal with the infinite series appearing blockwise in the product of $\infty \times \infty$ matrices and (2.5). We can justify them by virtue of the following simple fact: For any sequence $f_n \in \mathcal{gl}(r, \mathbb{C}[[x]])$ with $\text{ord}_x f_n \geq n$, the infinite series $\sum_{n=0}^{\infty} f_n$ is well defined as an element of $\mathcal{gl}(r, \mathbb{C}[[x]])$.)

Secondly, let us derive (1.5) and (1.14). To derive them, we have only to show that there is an $\infty \times \infty$ matrix $\tilde{C} = (\tilde{C}_{ij})_{i,j < 0}$ with $\tilde{C}_{ij} \in \mathcal{gl}(r, \mathbb{C}[[x]])$ such that

$$A\tilde{\xi} = \tilde{\xi}\tilde{C}, \quad (2.7)$$

$$\text{ord}_x \tilde{C}_{ij} \geq j - i - 1 (i \leq j - 1). \quad (2.8)$$

In fact, from (2.1) we have

$$(-A\partial_y + \partial_{\bar{z}})\tilde{\xi} = 0, \quad (A\partial_z + \partial_{\bar{y}})\tilde{\xi} = 0. \quad (2.9)$$

and applying Remark (iii) at the end of Sect. 1 to the transformation associated with $\tilde{\xi} \rightarrow \tilde{\xi}\tilde{\xi}_{(-)}^{-1}$, we can derive (1.5) and (1.14) from (2.7) and (2.9). The manipulation of the $\infty \times \infty$ matrices appearing in these discussions is justified by (2.4), (2.6) and (2.8).

To seek for such a matrix \tilde{C} , recall the correspondence $W^{(0)} \leftrightarrow \xi^{(0)}$ mentioned in Remark (ii) at the end of Sect. 1. From formula (1.9) applied to $\xi^{(0)}$, we have

$$\begin{aligned} \xi^{(0)} &= (W_{i-j}^{(0)*})_{i \in \mathbb{Z}, j < 0} (W_{i-j}^{(0)})_{i, j < 0}, \\ A_k \xi^{(0)} &= (W_{i-j}^{(0)*})_{i \in \mathbb{Z}, j < 0} A_{(-)}^k (W_{i-j}^{(0)})_{i, j < 0} \quad \text{for } k \geq 1, \end{aligned}$$

where $A_{(-)} = (\delta_{i+1, j} 1)_{i, j < 0}$. Then by virtue of the Leibniz formula and the above formulas,

$$\begin{aligned} (\bar{z}A\partial_y - \bar{y}A\partial_z)^k \xi^{(0)} &= \sum_{m=0}^k \binom{k}{m} [(\bar{z}A\partial_y - \bar{y}A\partial_z)^m (W_{i-j}^{(0)*})_{i \in \mathbb{Z}, j < 0}] \\ &\quad \cdot [(\bar{z}A_{(-)}\partial_y - \bar{y}A_{(-)}\partial_z)^{k-m} (W_{i-j}^{(0)})_{i, j < 0}] \end{aligned}$$

for $k \geq 1$, so that

$$\tilde{\xi} = \tilde{\Xi}\tilde{H}, \quad (2.10)$$

where

$$\begin{aligned} \tilde{\Xi} &= (\tilde{\Xi}_{ij})_{i \in \mathbb{Z}, j < 0} = \exp(\bar{z}A\partial_y - \bar{y}A\partial_z) (W_{i-j}^{(0)*})_{i \in \mathbb{Z}, j < 0}, \quad \tilde{\Xi}_{ij} \in \mathcal{gl}(r, \mathbb{C}[[x]]), \\ \tilde{H} &= (\tilde{H}_{ij})_{i, j < 0} = \exp(\bar{z}A_{(-)}\partial_y - \bar{y}A_{(-)}\partial_z) (W_{i-j}^{(0)})_{i, j < 0}, \quad \tilde{H}_{ij} \in \mathcal{gl}(r, \mathbb{C}[[x]]). \end{aligned}$$

Since $A\tilde{\Xi} = \tilde{\Xi}A_{(-)}$, as an immediate corollary we have

$$A\tilde{\xi} = \tilde{\Xi}A_{(-)}\tilde{H}. \quad (2.11)$$

Equation (2.11) shows that we may take $\tilde{C} = \tilde{H}^{-1}A_{(-)}\tilde{H}$, provided that the

product $\tilde{H}^{-1}A_{(-)}\tilde{H}$ can be defined as an $\infty \times \infty$ matrix consisting of $r \times r$ blocks $\in \mathcal{gl}(r, \mathbb{C}[[x]])$ such that ord_x [the (ij) block of $\tilde{H}^{-1}A_{(-)}\tilde{H} \geq j - i - 1 (i \leq j - i)$. We can actually justify this in the same way as the justification of $\tilde{\xi}_{(+)}\tilde{\xi}_{(-)}^{-1}$. First, note that

$$\begin{aligned} \tilde{H} &= \mathbb{1} + \tilde{H}, \\ \tilde{H} &= (\tilde{H}_{ij})_{i,j < 0} = \exp(\bar{z}A_{(-)}\partial_y - \bar{y}A_{(-)}\partial_z)(W_{i-j}^{(0)} - \delta_{ij}1_r)_{i,j < 0}. \end{aligned}$$

It is easy to show that $\tilde{H}_{ij} \in \mathcal{gl}(r, \mathbb{C}[[x]])$ and $\text{ord}_x \tilde{H}_{ij} \geq j - i + 1 (i \leq j + 1)$. Hence we can define the k^{th} power \tilde{H} as an $\infty \times \infty$ matrix with blocks in $\mathcal{gl}(r, \mathbb{C}[[x]])$ such that ord_x [the (ij) block of $\tilde{H}^k \geq j - i + k (i \leq j + k)$. Hence the inverse of \tilde{H} is given by the Neumann series

$$\tilde{H}^{-1} = \mathbb{1} + \sum_{k=1}^{\infty} (-\tilde{H})^k,$$

which consequently satisfies the conditions that ord_x [the (ij) block of $\tilde{H}^{-1} \geq j - i (i \leq j)$. From these observations we can show immediately that the product $\tilde{H}^{-1}A_{(-)}\tilde{H}$ can be defined and satisfies all the required conditions. Thus we can take $\tilde{C} = \tilde{H}^{-1}A_{(-)}\tilde{H}$.

Finally the initial condition follows immediately from the construction. This completes the proof of Theorem 6.

2.3.

We here prove Theorem 7.

First, we can show from the assumption that the characteristic matrices $G_{ij}^{(0)}$ of $W^{(0)}$ are matrix-valued holomorphic functions of (y, z) defined in a neighborhood U of $(y, z) = (0, 0)$ with the estimates

$$|\partial_y^k \partial_z^m G_{ij}^{(0)}| \leq M a^{k+m} b^{i+j} k! m! \quad \text{for } i, j \geq 1, k, m \geq 0 \text{ and } (y, z) \in U, \quad (2.12)$$

where M, a and b are positive constants, and the sign $|\cdot|$ denotes the norm of $r \times r$ matrices, $|(a_{\alpha\beta})_{1 \leq \alpha, \beta \leq r}| = \max_{1 \leq \alpha, \beta \leq r} |a_{\alpha\beta}|$. In fact, from formula (1.11) applied to $G_{ij}^{(0)}$ we have

$$\sum_{i,j=1}^{\infty} G_{ij}^{(0)} \mu^{-1} \lambda^{-j} = (W^{(0)}(y, z, \mu)^{-1} W^{(0)}(y, z, \lambda) - 1_r) / (\lambda - \mu),$$

while the assumption together with the formula $W^{(0)}(y, z, \infty) = 1_r$, implies that the right-hand side of the above formula is a matrix-valued holomorphic function of (y, z, λ, μ) defined in a neighborhood of $(y, z, \lambda, \mu) = (0, 0, \infty, \infty)$ in $\mathbb{C}^2 \times \mathbb{P}^1 \times \mathbb{P}^1$. Hence (2.12) follows by virtue of Cauchy's inequality.

We can estimate the blocks of $\tilde{\xi}$ as follows. Proposition 4 shows that $\xi_{ij}^{(0)} = G_{i+1, -j}^{(0)} (i \geq 0, j < 0)$, so that from (2.3)

$$\tilde{\xi}_{ij} = \sum_{k=\max(0, -i)}^{\infty} (\bar{z}\partial_y - \bar{y}\partial_z)^k G_{i+k+1, -j}^{(0)} / k!$$

On the other hand, applying (2.12) to the right-hand side of

$$(\bar{z}\partial_y - \bar{y}\partial_z)^k G_{i+k+1, -j}^{(0)} = \sum_{m=0}^k \binom{k}{m} \bar{z}^m (-\bar{y})^{k-m} \partial_y^m \partial_z^{k-m} G_{i+k+1, -j}^{(0)},$$

we have

$$|(\bar{z}\partial_y - \bar{y}\partial_z)^k G_{i+k+1, -j}^{(0)}| \leq Ma^k b^{i+k+1-j} (|\bar{y}| + |\bar{z}|)^k k!$$

Hence $\tilde{\xi}_{ij}$ is a matrix-valued holomorphic function of x in the domain $2ab(|\bar{y}| + |\bar{z}|) < 1, (y, z) \in U$ in \mathbb{C}^4 with the estimate

$$\begin{aligned} |\tilde{\xi}_{ij}| &\leq \sum_{k=\max\{0, -i\}}^{\infty} Ma^k b^{i+k+1-j} (|\bar{y}| + |\bar{z}|)^k \\ &= Ma^{\max\{0, -1\}} b^{i+1-j+\max\{0, -i\}} (|\bar{y}| + |\bar{z}|)^{\max\{0, -i\}} / (1 - ab(|\bar{y}| + |\bar{z}|)) \\ &\leq 2Mb^{i+1-j}, \quad \text{if } i \geq 0, \end{aligned} \tag{2.13a}$$

$$\leq 2Ma^{-i} b^{1-j} (|\bar{y}| + |\bar{z}|)^{-i}, \quad \text{if } i < 0. \tag{2.13b}$$

Now, the uniform convergence of the Neumann series in (2.5) can be checked as follows. First, if $2ab(|\bar{y}| + |\bar{z}|) < 1$ and $(y, z) \in U$, we can prove the following estimates for $k \geq 1$ by induction:

$$\begin{aligned} &|\text{the } (ij) \text{ block of } \tilde{\xi}_{(-)}^k| \\ &\leq 2^{2k-1} M^k r^{k-1} a^{k-1-i} b^{2k-1-j} (|\bar{y}| + |\bar{z}|)^{k-1-i} \text{ for } i, j < 0. \end{aligned} \tag{2.14}$$

When $k = 1$, this is nothing but (2.13b). Suppose that (2.14) is valid for k . Then from (2.13b) and the assumption that $2ab(|\bar{y}| + |\bar{z}|) < 1$,

$$\begin{aligned} &|\text{the } (ij) \text{ block of } \tilde{\xi}_{(-)}^{k+1}| \\ &\leq r \sum_{m < 0} |\text{the } (im) \text{ block of } \tilde{\xi}_{(-)}^k| |\tilde{\xi}_{mj}| \\ &\leq 2^{2k} M^{k+1} r^k a^{k-i} b^{2k+1-j} (|\bar{y}| + |\bar{z}|)^{k-i} \sum_{m < 0} (ab(|\bar{y}| + |\bar{z}|))^{-m-1} \\ &\leq 2^{2k+1} M^{k+1} r^k a^{k-i} b^{2k+1-j} (|\bar{y}| + |\bar{z}|)^{k-i}. \end{aligned}$$

This completes the induction. On the other hand, summing up the right-hand side of (2.14) over $k \geq 1$, we get an infinite series which converges uniformly in the domain $V = \{x = (y, z, \bar{y}, \bar{z}) \in \mathbb{C}^4; 2ab(|\bar{y}| + |\bar{z}|) < 1, 8Mrab^2(|\bar{y}| + |\bar{z}|) < 1, (y, z) \in U\}$ and coincides with $2Ma^{-i} b^{1-j} (|\bar{y}| + |\bar{z}|)^{-i} / (1 - 4Mrab^2(|\bar{y}| + |\bar{z}|))$. Hence the Neumann series in (2.5) converges blockwise uniformly in V , and defines a matrix-valued holomorphic function in V with the estimates

$$|\text{the } (ij) \text{ block of } \tilde{\xi}_{(-)}^{-1}| \leq \delta_{ij} + 4Ma^{-i} b^{1-j} (|\bar{y}| + |\bar{z}|)^{-i}. \tag{2.15}$$

Finally, (2.13a) and (2.15) show that $\tilde{\xi}_{(+)} \tilde{\xi}_{(-)}^{-1} = (G_{i+1, -j})_{i \geq 0, j < 0}$ defines a matrix-valued holomorphic function of x in V with the estimates

$$|G_{i+1, -j}| \leq r \sum_{m < 0} |\tilde{\xi}_{im}| |\text{the } (mj) \text{ block of } \tilde{\xi}_{(-)}^{-1}|$$

$$\begin{aligned} &\leq 2Mr b^{i+1-j} + 8M^2 rab^{i+3-j}(|\bar{y}| + |\bar{z}|) \sum_{m < 0} (ab(|\bar{y}| + |\bar{z}|))^{-m-1} \\ &\leq (2Mr + 16M^2 rab^2(|\bar{y}| + |\bar{z}|)) b^{i+1-j}. \end{aligned} \tag{2.16}$$

This shows, in particular, that $W = 1_r - \sum_{j=1}^{\infty} G_{ij} \lambda^{-j}$ converges uniformly to define a matrix-valued holomorphic function of (x, λ) in a neighborhood of $(x, \lambda) = (0, \infty)$ in $\mathbb{C}^4 \times \mathbb{P}^1$. (As a matter of fact, (2.16) also implies the convergence of $\sum_{i,j=1}^{\infty} G_{ij} \mu^{-i} \lambda^{-j}$ in a neighborhood of $(x, \lambda, \mu) = (0, \infty, \infty)$ in $\mathbb{C}^4 \times \mathbb{P}^1 \times \mathbb{P}^1$.) This completes the proof of Theorem 7.

3. Construction of Transformation

In this section we discuss transformations acting on the space of solutions to (0.3).

3.1

We first consider the formal aspects.

As the data for a transformation, let us consider an $\infty \times \infty$ matrix $\mathbb{P} = (P_{ij})_{i,j \in \mathbb{Z}}$ (with $P_{ij} \in \mathcal{g}^{\ell}(r, \mathbb{C}[[x]])$) satisfying the following conditions

$$[A, \mathbb{P}] = 0, [-A\partial_y + \partial_z, \mathbb{P}] = 0, [A\partial_z + \partial_y, \mathbb{P}] = 0. \tag{3.1}$$

(If necessary, we may suppose that the blocks P_{ij} depend on some other variables parametrizing \mathbb{P} . Later we shall encounter such a case.) Note that the first condition in (3.1) means that \mathbb{P} has the form

$$\mathbb{P} = (P_{j-i})_{i,j \in \mathbb{Z}}, P_j \in \mathcal{g}^{\ell}(r, \mathbb{C}[[x]]), \tag{3.2}$$

and that in terms of the generating function $P = \sum_{j \in \mathbb{Z}} P_j \lambda^j$ the remaining ones read $(-\lambda\partial_y + \partial_z)P = 0, (\lambda\partial_z + \partial_y)P = 0$.

The transformation associated with \mathbb{P} is constructed as follows. Let $\xi = (\xi_{ij})_{i \in \mathbb{Z}, j < 0}$ (with $\xi_{ij} \in \mathcal{g}^{\ell}(r, \mathbb{C}[[x]])$) be a solution to (0.3). Then $\mathbb{P}\xi$ satisfies (1.5) and (1.14), but not (1.6). In order to achieve (1.5), we perform the procedure mentioned in Remark (i) at the end of Sect. 1. Namely, dividing the matrices ξ and \mathbb{P} into the blocks of the form

$$\begin{aligned} \xi &= \begin{pmatrix} \mathbb{1} \\ \mathbb{G} \end{pmatrix} \text{ with } \mathbb{1} = (\delta_{ij} 1_r)_{i,j < 0} \text{ and } \mathbb{G} = (\xi_{ij})_{i \geq 0, j < 0}, \\ \mathbb{P} &= \begin{pmatrix} \mathbb{P}_1 & \mathbb{P}_4 \\ \mathbb{P}_2 & \mathbb{P}_3 \end{pmatrix} \text{ with } \mathbb{P}_1 = (P_{j-i})_{i,j < 0}, \mathbb{P}_4 = (P_{j-i})_{i < 0, j \geq 0}, \\ &\quad \mathbb{P}_2 = (P_{j-i})_{i \geq 0, j < 0} \text{ and } \mathbb{P}_3 = (P_{j-i})_{i,j \geq 0}, \end{aligned} \tag{3.3}$$

and multiplying $\mathbb{P}\xi$ by $(\mathbb{P}_1 + \mathbb{P}_4 \mathbb{G})^{-1}$ on the assumption that $\mathbb{P}_1 + \mathbb{P}_4 \mathbb{G}$ is invertible, we finally obtain the matrix

$$\mathbb{P} \circ \xi = \mathbb{P}\xi (\mathbb{P}_1 + \mathbb{P}_4 \mathbb{G})^{-1} = \begin{pmatrix} \mathbb{1} \\ \mathbb{P} \circ \mathbb{G} \end{pmatrix}, \tag{3.4}$$

which satisfies (0.3), where

$$\mathbb{P} \circ \mathbb{G} = (\mathbb{P}_2 + \mathbb{P}_3 \mathbb{G})(\mathbb{P}_1 + \mathbb{P}_4 \mathbb{G})^{-1}. \quad (3.5)$$

Thus we have shown that any matrix \mathbb{P} satisfying (3.1) induces the transformation $\xi \rightarrow \mathbb{P} \circ \xi$ in the space of solutions to (0.3) (as far as we can justify the definition of $\mathbb{P} \circ \xi$.) Also, it is remarkable that in terms of the matrix \mathbb{G} this transformation have the form of a linear fractional transformation $\mathbb{G} \rightarrow \mathbb{P} \circ \mathbb{G}$ of $\infty \times \infty$ matrices. This fact seems to suggest some relation to the discussions of Hauser and Ernst [19, Sect. V].

It is also possible, at least formally, to derive the associated infinitesimal transformation. To see this, we set

$$\mathbb{P} = \mathbb{1} + \varepsilon \mathbb{P}, \quad \mathbb{P} = \begin{pmatrix} \mathbb{P}_1 & \mathbb{P}_4 \\ \mathbb{P}_2 & \mathbb{P}_3 \end{pmatrix}, \quad (3.6)$$

and calculate $\mathbb{P} \circ \mathbb{G}$ modulo ε^2 , i.e. to the first order with respect to ε . Since $\mathbb{P}_1 = \mathbb{1} + \varepsilon \mathbb{P}_1$ and $\mathbb{P}_3 = \mathbb{1} + \varepsilon \mathbb{P}_3$, they are to be considered invertible with $\mathbb{P}_1^{-1} \equiv \mathbb{1} - \varepsilon \mathbb{P}_1$ and $\mathbb{P}_3^{-1} \equiv \mathbb{1} - \varepsilon \mathbb{P}_3 \pmod{\varepsilon^2}$. Hence $\mathbb{1} + \mathbb{P}_4 \mathbb{G} \mathbb{P}_1^{-1} \equiv \mathbb{1} + \varepsilon \mathbb{P}_4 \mathbb{G} (\mathbb{1} - \varepsilon \mathbb{P}_1) \equiv \mathbb{1} + \varepsilon \mathbb{P}_4 \mathbb{G} \pmod{\varepsilon^2}$, so that

$$(\mathbb{1} + \mathbb{P}_4 \mathbb{G} \mathbb{P}_1^{-1})^{-1} \equiv \mathbb{1} - \varepsilon \mathbb{P}_4 \mathbb{G} \pmod{\varepsilon^2}.$$

Similarly,

$$\begin{aligned} \mathbb{P}_2 \mathbb{P}_1^{-1} + \mathbb{P}_3 \mathbb{G} \mathbb{P}_1^{-1} &\equiv \varepsilon \mathbb{P}_2 (\mathbb{1} - \varepsilon \mathbb{P}_1) + (\mathbb{1} + \varepsilon \mathbb{P}_3) \mathbb{G} (\mathbb{1} - \varepsilon \mathbb{P}_1) \\ &\equiv \mathbb{G} + \varepsilon (\mathbb{P}_2 + \mathbb{P}_3 \mathbb{G} - \mathbb{G} \mathbb{P}_1) \pmod{\varepsilon^2}. \end{aligned}$$

Hence we obtain the following representation of the infinitesimal transformation with generator \mathbb{P} :

$$\begin{aligned} (\mathbb{1} + \varepsilon \mathbb{P}) \circ \mathbb{G} &= (\mathbb{P}_2 \mathbb{P}_1^{-1} + \mathbb{P}_3 \mathbb{G} \mathbb{P}_1^{-1}) (\mathbb{1} + \mathbb{P}_4 \mathbb{G} \mathbb{P}_1^{-1})^{-1} \\ &\equiv \mathbb{G} + \varepsilon (\mathbb{P}_2 + \mathbb{P}_3 \mathbb{G} - \mathbb{G} \mathbb{P}_1 - \mathbb{G} \mathbb{P}_4 \mathbb{G}) \pmod{\varepsilon^2}. \end{aligned}$$

This representation is very similar to those presented in [12], though they are basically different from ours: The characteristic matrices used in [12] for representing the infinitesimal transformations are associated with W_+ of the pair (W_+, W_-) in the Riemann–Hilbert problem, while our W corresponds to W_- . (Here the notations are the same as those used in Sect. 4).

We can also describe the effect of the above transformations on the initial values. As can be checked easily, $\xi \rightarrow \mathbb{P} \circ \xi$ induces the transformation $\xi^{(0)} \rightarrow \mathbb{P}^{(0)} \circ \xi^{(0)}$, where $\xi^{(0)} = \xi|_{y=z=0}$, $\mathbb{P}^{(0)} = \mathbb{P}|_{y=z=0}$, and $\mathbb{P}^{(0)} \circ \xi^{(0)}$ is defined in the same way as the definition of $\mathbb{P} \circ \xi$, and the latter transformation uniquely determines the former. The second and third conditions in (3.1) imply that \mathbb{P} is recovered from $\mathbb{P}^{(0)}$ by

$$\mathbb{P} = \exp(\bar{z} A \partial_y - \bar{y} A \partial_z) \mathbb{P}^{(0)},$$

and the only condition to be imposed on $\mathbb{P}^{(0)}$ is that $[A, \mathbb{P}^{(0)}] = 0$, so that $\mathbb{P}^{(0)}$ is written in the form $\mathbb{P}^{(0)} = (P_j^{(0)})_{i,j \in \mathbb{Z}}$ with $P_j^{(0)} \in \mathfrak{gl}(r, \mathbb{C}[[y, z]])$.

3.2

We here consider, as examples of the cases where the above arguments are rigorously justified, two important classes of transformations. We call them the transformations of type I and type II respectively. They are defined by the following conditions imposed on $\mathbb{P}^{(0)} = (P_{j-i}^{(0)})_{i,j \in \mathbb{Z}}$.

Type I: $P_j^{(0)} = 0$ for $j > 0$ and $P_0^{(0)} = 1_r$.

Type II: $P_j^{(0)} = 0$ for $j < 0$ and $P_0^{(0)}$ is invertible.

In the case of the transformations of type I, the definition of $\mathbb{P} \circ \xi$ is justified as follows. From (3.5) and the conditions imposed on $\mathbb{P}^{(0)}$, we have

$$\text{ord}_x(P_j - \delta_{j0}1_r) \geq j + 1 \quad \text{for } j \geq 0.$$

Bearing this fact in mind, we can show in the same way as the discussions developed in Sect. 2 that the product $(\mathbb{P}_2 + \mathbb{P}_3 \mathbb{G}) \cdot (\mathbb{P}_1 + \mathbb{P}_4 \mathbb{G})^{-1}$ can be defined as an $\infty \times \infty$ matrix consisting of $r \times r$ blocks $\in \mathcal{gl}(r, \mathbb{C}[[x]])$, where the inverse of $\mathbb{P}_1 + \mathbb{P}_4 \mathbb{G}$ is given by the Neumann series

$$(\mathbb{P}_1 + \mathbb{P}_4 \mathbb{G})^{-1} = 1 + \sum_{k=1}^{\infty} (1 - \mathbb{P}_1 - \mathbb{P}_4 \mathbb{G})^k.$$

The effect of the transformations of type I on the space of initial values is extremely simple. To see this, we recall the correspondence $W^{(0)} \rightarrow \xi^{(0)}$ mentioned in Remark (ii) at the end of Sect. 1.

Proposition 9. *The transformations of type I induce the transformations $W^{(0)} \rightarrow W^{(0)}P^{(0)-1}$, where $P^{(0)} = \sum_{j \leq 0} P_j^{(0)} \lambda^j$.*

Proof. From formula (1.9) applied to $\xi^{(0)}$ we have

$$\xi^{(0)} = (W_{i-j}^{(0)*})_{i \in \mathbb{Z}, j < 0} (W_{i-j}^{(0)})_{i, j < 0},$$

so that

$$\begin{aligned} \mathbb{P}^{(0)} \circ \xi^{(0)} &= \mathbb{P}^{(0)} \xi^{(0)} (P_{j-i}^{(0)})_{i,j < 0}^{-1} \\ &= (P_{j-i}^{(0)})_{i,j \in \mathbb{Z}} (W_{i-j}^{(0)})_{i,j \in \mathbb{Z}}^{-1} (\delta_{ij}1_r)_{i \in \mathbb{Z}, j < 0} (W_{j-i}^{(0)})_{i,j < 0} (P_{j-i}^{(0)})_{i,j < 0}^{-1} \\ &= (\tilde{W}_{i-j}^{(0)*})_{i \in \mathbb{Z}, j < 0} (\tilde{W}_{i-j}^{(0)})_{i,j < 0}, \end{aligned}$$

where $\tilde{W}_j^{(0)} = \tilde{W}_j^{(0)*} = 0$ for $j < 0$, and $\tilde{W}_j^{(0)}$, $\tilde{W}_j^{(0)*}$, $j \geq 0$, denote respectively the coefficients of $W^{(0)}P^{(0)-1}$ and $P^{(0)}W^{(0)-1}$, i.e. $W^{(0)}P^{(0)-1} = \sum_{j=0}^{\infty} \tilde{W}_j^{(0)} \lambda^{-j}$,

$P^{(0)}W^{(0)-1} = \sum_{j=0}^{\infty} \tilde{W}_j^{(0)*} \lambda^{-j}$. This proves the proposition.

This proposition shows that the totality of the transformations of type I constitutes a group isomorphic to the group $\{P^{(0)}; P^{(0)} = \sum_{j \leq 0} P_j^{(0)} \lambda^j, P_j^{(0)} \in \mathcal{gl}(r, \mathbb{C}[[x]]), P_0^{(0)} = 1_r\}$, whose group structure is defined by the usual product $P^{(0)}, \tilde{P}^{(0)} \rightarrow P^{(0)}\tilde{P}^{(0)}$. Note that this group acts transitively on the space of all the formal power series solutions to (0.3).

In the case of the transformations of type II, we may use the matrix-valued variables $P_{j\rho\sigma}^{(0)}$, $j, \rho, \sigma \geq 0$, defined by the expansion

$$P_j^{(0)} = \sum_{\rho, \sigma=0}^{\infty} P_{j\rho\sigma}^{(0)} y^\rho z^\sigma$$

to parametrize the matrix \mathbb{P} in the form

$$\mathbb{P} = \sum_{j, \rho, \sigma=0}^{\infty} P_{j\rho\sigma}^{(0)} A^j (y + A\bar{z})^\rho (z - A\bar{y})^\sigma,$$

where $P_{j\rho\sigma}^{(0)}$ is identified with the $\infty \times \infty$ matrix $(P_{j\rho\sigma}^{(0)} \delta_{ij})_{i, j \in \mathbb{Z}}$. Then we can justify the definition of $\mathbb{P} \circ \xi$ in a suitable noncommutative algebra of formal power series of $x, P_{j\rho\sigma}^{(0)}$ and $P_{000}^{(0)-1}$, extending the algebra $\mathcal{A}[[x]]$,

We, however, omit the detail of these discussions, since introducing the matrix-valued formal variables makes the algebraic setting much too complicated. Instead, let us consider below a simplified case. Even in that case we can see the essence of the discussions to be developed in the general case.

What we here consider is the case where $P_j^{(0)}, j \geq 0$, are diagonal matrices. In this case we can introduce new independent variables $t = (t_{n\rho\sigma}^{(\alpha)})_{1 \leq \alpha \leq r, n, \rho, \sigma \geq 0}$ to parametrize \mathbb{P} in the form

$$\mathbb{P} = \exp \left(\sum_{\alpha=1}^r \sum_{n, \rho, \sigma=0}^{\infty} t_{n\rho\sigma}^{(\alpha)} E_\alpha A^n (y + A\bar{z})^\rho (z + A\bar{y})^\sigma \right), \quad (3.6)$$

where $E_\alpha, 1 \leq \alpha \leq r$, denote the $\infty \times \infty$ matrices

$$E_\alpha = ((\delta_{\alpha\beta} \delta_{\alpha\gamma})_{1 \leq \beta, \gamma \leq r} \delta_{ij})_{i, j \in \mathbb{Z}},$$

which from now on we often identify with the $r \times r$ matrices $(\delta_{\alpha\beta} \delta_{\alpha\gamma})_{1 \leq \beta, \gamma \leq r}$. Note that $E_\alpha, 1 \leq \alpha \leq r$, and A commute each other and satisfy the formula $\sum_{\alpha=1}^r E_\alpha = \mathbb{1}$.

Let us consider the justification of the definition of $\mathbb{P} \circ \xi$, which from now on is denoted by $\xi(x, t)$ (for the later use), i.e.

$$\xi(x, t) = \mathbb{P} \circ \xi = \left(\begin{array}{c} (\delta_{ij} 1)_{i, j < 0} \\ (G_{i+1, -j}(x, t))_{i \geq 0, j < 0} \end{array} \right). \quad (3.7)$$

Here $G_{ij}(x, t), i \geq 1, j \geq 1$, denote the corresponding characteristic matrices (cf. Proposition 4). According to the construction of transformations presented before, $\xi(x, t)$ is to be defined by

$$\xi(x, t) = \tilde{\xi}(x, t) \tilde{\xi}(x, t)_{(-)}^{-1}, \quad (3.8)$$

where

$$\begin{aligned} \tilde{\xi}(x, t) &= (\tilde{\xi}_{ij}(x, t))_{i \in \mathbb{Z}, j < 0} = \mathbb{P} \xi, \\ \tilde{\xi}(x, t)_{(-)} &= (\tilde{\xi}_{ij}(x, t))_{i, j < 0}. \end{aligned} \quad (3.9)$$

We now introduce an algebra $\mathbb{C}[[x, t]]$ of formal power series of x and t by

$$\mathbb{C}[[x, t]] = \left\{ f; f = \sum_{n=0}^{\infty} f_n, f_n \in \mathbb{C}[[x, t_{000}^{(\alpha)} (1 \leq \alpha \leq r)]], \text{ord}_i f_n \geq n(n \geq 0) \right\},$$

where the notation $\text{ord}_t f_n \geq n$ means that f_n is a linear combination, with coefficients in $\mathbb{C}[[x, t_{000}^{(\alpha)} (1 \leq \alpha \leq r)]]$, of only the monomials of the form $t_{n_1 \rho_1 \sigma_1}^{(\alpha_1)} t_{n_2 \rho_2 \sigma_2}^{(\alpha_2)} \dots$ with $n_1 + \rho_1 + \sigma_1 + n_2 + \rho_2 + \sigma_2 + \dots \geq n$. Then we can show in the same way as the discussions in Sect. 2, using ord_t instead of ord_x , that the inverse $\tilde{\xi}(x, t)_{(-)}^{-1}$ and the product $\tilde{\xi}(x, t)\tilde{\xi}(x, t)_{(-)}^{-1}$ can be defined as $\infty \times \infty$ matrices with blocks in $\mathcal{gl}(r, \mathbb{C}[[x, t]])$, and that ord_t [the (ij) block of $\tilde{\xi}(x, t)_{(-)}^{\pm 1}$] $\geq j - i$ ($i \leq j$). In particular, $G_{ij}(x, t) \in \mathcal{gl}(r, \mathbb{C}[[x, t]])$.

3.2

The transformation $\xi \rightarrow \tilde{\xi}(x, t)$ turns out to have a remarkable feature. Namely, this induces a hierarchy of nonlinear differential equations which are similar to the AKNS systems [20] and consistently coupled with the self-dual Yang–Mills equations, as we shall see below.

First, in terms of $\tilde{\xi}(x, t)$ we have

Proposition 10. $\tilde{\xi}(x, t)$ satisfies the equations

$$\begin{aligned} (-A\partial_y + \partial_{\bar{z}})\tilde{\xi}(x, t) + \tilde{\xi}(x, t)A(x, t) &= 0, \\ (A\partial_z + \partial_{\bar{y}})\tilde{\xi}(x, t) + \tilde{\xi}(x, t)B(x, t) &= 0, \\ A\tilde{\xi}(x, t) &= \tilde{\xi}(x, t)C(x, t), \end{aligned} \quad (3.10)$$

and

$$\partial \tilde{\xi}(x, t) / \partial t_{n\rho\sigma}^{(\alpha)} = E_\alpha A^n (y + A\bar{z})^\rho (z - A\bar{y})^\sigma (x, t) - \tilde{\xi}(x, t) C_{n\rho\sigma}^{(\alpha)}(x, t) \quad (3.11)$$

for $1 \leq \alpha \leq r$ and $n, \rho, \sigma \geq 0$, where

$$\begin{aligned} A(x, t) &= \begin{pmatrix} 0 \\ (\partial_y G_{1, -j}(x, t))_{j < 0} \end{pmatrix}, \quad B(x, t) = \begin{pmatrix} 0 \\ (-\partial_z G_{1, -j}(x, t))_{j < 0} \end{pmatrix}, \\ C(x, t) &= \begin{pmatrix} (\delta_{i+1, j} 1_r)_{i < -1, j < 0} \\ (G_{1, -j}(x, t))_{j < 0} \end{pmatrix}, \\ C_{n\rho\sigma}^{(\alpha)}(x, t) &= E_\alpha C(x, t)^n (y + C(x, t)\bar{z})^\rho (z - C(x, t)\bar{y})^\sigma. \end{aligned} \quad (3.12)$$

Proof. Note that by virtue of (3.6) and (3.9) $\tilde{\xi}(x, t)$ satisfies

$$\begin{aligned} (-A\partial_y + \partial_{\bar{z}})\tilde{\xi}(x, t) + \tilde{\xi}(x, t)A &= 0, \quad (A\partial_z + \partial_{\bar{y}})\tilde{\xi}(x, t) + \tilde{\xi}(x, t)B = 0, \\ A\tilde{\xi}(x, t) = \tilde{\xi}(x, t)C, \quad \partial \tilde{\xi}(x, t) / \partial t_{n\rho\sigma}^{(\alpha)} &= E_\alpha A^n (y + A\bar{z})^\rho (z - A\bar{y})^\sigma \tilde{\xi}(x, t), \end{aligned}$$

where A, B and C are the same as those given by (1.7) and (1.15). The argument developed in Remark (iii) at the end of Sect. 1 can be easily generalized to conclude that the above equations satisfied by $\tilde{\xi}(x, t)$ imply (3.10)–(3.12). This proves the proposition.

Next, let $W(x, t, \lambda) = \sum_{j=0}^{\infty} W_j(x, t) \lambda^{-j}$ denote the formal power series corresponding to $\tilde{\xi}(x, t)$ via the correspondence established in Sect. 1, i.e. $W_0(x, t) = 1$, and $W_j(x, t) = -G_{1j}(x, t)$ for $j \geq 1$. In terms of $W(x, t, \lambda)$ we have

Proposition 11. $W(x, t, \lambda)$ satisfies the equations

$$\begin{aligned} (-\lambda\partial_y + \partial_{\bar{z}} + A_{\bar{z}}(x, t))W(x, t, \lambda) &= 0, \\ (\lambda\partial_z + \partial_{\bar{y}} + A_{\bar{y}}(x, t))W(x, t, \lambda) &= 0, \end{aligned} \quad (3.13)$$

and

$$\partial W(x, t, \lambda)/\partial t_{n\rho\sigma}^{(\alpha)} = B_{n\rho\sigma}^{(\alpha)}(x, t, \lambda)W(x, t, \lambda) - W(x, t, \lambda)E_\alpha \lambda^n (y + \lambda\bar{z})^\rho (z - \lambda\bar{y})^\sigma, \quad (3.14)$$

for $1 \leq \alpha \leq r$ and $n, \rho, \sigma \geq 0$, where

$$\begin{aligned} A_{\bar{y}}(x, t) &= -\partial_z W_1(x, t), \quad A_z(x, t) = \partial_y W_1(x, t), \\ B_{n00}^{(\alpha)}(x, t, \lambda) &= \sum_{m=0}^n \sum_{k=0}^m W_k(x, t) E_\alpha W_{m-k}^*(x, t) \lambda^{n-m}, \\ B_{n\rho\sigma}^{(\alpha)}(x, t, \lambda) &= \sum_{\rho'=0}^{\rho} \sum_{\sigma'=0}^{\sigma} \binom{\rho}{\rho'} \binom{\sigma}{\sigma'} y^{\rho-\rho'} \bar{z}^{\rho'} z^{\sigma-\sigma'} (-\bar{y})^{\sigma'} \\ &\quad \cdot B_{n+\rho'+\sigma'00}^{(\alpha)}(x, t, \lambda), \end{aligned} \quad (3.15)$$

and $W_j^*(x, t), j \geq 0$, denote the coefficients of $W(x, t, \lambda)^{-1}$, i.e.

$$W(x, t, \lambda)^{-1} = \sum_{j=0}^{\infty} W_j(x, t) \lambda^{-j}.$$

As a corollary of the above propositions, we find, after some calculations, that

$$\begin{aligned} [-\lambda\partial_y + \partial_{\bar{z}} + A_{\bar{z}}(x, t), \lambda\partial_z + \partial_{\bar{y}} + A_{\bar{y}}(x, t)] &= 0, \\ [-\lambda\partial_y + \partial_{\bar{z}} + A_{\bar{z}}(x, t), \partial/\partial t_{n\rho\sigma}^{(\alpha)} - B_{n\rho\sigma}^{(\alpha)}(x, t, \lambda)] &= 0, \\ [\lambda\partial_z + \partial_{\bar{y}} + A_{\bar{y}}(x, t), \partial/\partial t_{n\rho\sigma}^{(\alpha)} - B_{n\rho\sigma}^{(\alpha)}(x, t, \lambda)] &= 0, \\ [\partial/\partial t_{n\rho\sigma}^{(\alpha)} - B_{n\rho\sigma}^{(\alpha)}(x, t, \lambda), \partial/\partial t_{n'\rho'\sigma'}^{(\alpha')} - B_{n'\rho'\sigma'}^{(\alpha')}(x, t, \lambda)] &= 0, \end{aligned}$$

for $1 \leq \alpha, \alpha' \leq r$, and $n, n', \rho, \rho', \sigma, \sigma' \geq 0$. (3.16)

If we forget the dependence on x , the last equations in (3.16) are generalizations of the AKNS systems [17] to the rank r case. The whole system (3.16) provides a consistent coupling of them with the self-dual Yang–Mills equations.

Now, let us proceed to the proof of Proposition 11. We shall prove the equivalence of the equations imposed on $\xi(x, t)$ and those imposed on $W(x, t, \lambda)$. Since the equivalence of (3.10) and (3.13) has already been established in Sect. 1, it remains only to prove that of (3.11) and (3.14). We shall, at first, consider the case where $\rho = \sigma = 0$, for simplicity. The general case will be discussed later.

To begin with, we rewrite the equation for $\rho = \sigma = 0$ in (3.11) in terms of generating functions of $G_{ij}(x, t)$. First, comparing the “upper half” parts of both sides of the formula $A^n \xi(x, t) = \xi(x, t)C(x, t)^n$ (cf. Remark (iii) at the end of Sect. 1) and multiplying them by E_α , we have

$$C_{n00}^{(\alpha)}(x, t) = E_\alpha C(x, t)^n = \begin{pmatrix} (\delta_{i+n, j} E_\alpha)_{i < -n, j < 0} \\ (E_\alpha G_{i+n+1, -j}(x, t))_{-n \leq i < 0, j < 0} \end{pmatrix}.$$

By virtue of this formula, we can rewrite the equation for $\rho = \sigma = 0$ in (3.11) into

the equation

$$\partial G_{ij}(x, y)/\partial t_{n00}^{(\alpha)} = E_\alpha G_{i+n,j}(x, t) - G_{i+n,j}(x, t)E_\alpha - \sum_{m=1}^n G_{i,n-m+1}(x, t)E_\alpha G_{mj}(x, t).$$

Finally, we can once more rewrite them in terms of generating functions of $G_{ij}(x, t)$ with indeterminate variables λ and μ to obtain

$$\begin{aligned} \partial \sum_{i,j=1}^{\infty} G_{ij}(x, t)\mu^{-i}\lambda^{-j}/\partial t_{n00}^{(\alpha)} \\ = \mu^n E_\alpha \sum_{j=1}^{\infty} G_{ij}(x, t)\mu^{-i}\lambda^{-j} - \lambda^n \sum_{i,j=1}^{\infty} G_{ij}(x, t)\mu^{-i}\lambda^{-j}E_\alpha + \sum_{j=1}^n \sum_{i=1}^{\infty} G_{ij}(x, t)\mu^{-i}\lambda^{n-j}E_\alpha \\ - E_\alpha \sum_{i=1}^n \sum_{j=1}^{\infty} G_{ij}(x, t)\lambda^{-j}\mu^{n-i} - \sum_{m=1}^n \sum_{i=1}^{\infty} G_{i,n-m+1}(x, t)\lambda^{-i}E_\alpha \sum_{j=1}^{\infty} G_{mj}(x, t)\mu^{-j}. \end{aligned} \tag{3.17}$$

To calculate the right-hand side of (3.17) further, we prepare the following.

Lemma 12. *We have*

$$\begin{aligned} \sum_{i=1}^{\infty} G_{ij}(x, t)\mu^{-i} &= W(x, t, \mu)^{-1} \sum_{k=0}^{j-1} W_k(x, t)\mu^{j-k-1} - \mu^{j-1}1_r, \\ \sum_{j=1}^{\infty} G_{ij}(x, t)\lambda^{-j} &= \lambda^{i-1}1_r - \sum_{k=0}^{i-1} W_k^*(x, y)\lambda^{i-k-1}W(x, t, \lambda). \end{aligned}$$

Proof. The calculation in the proof of Proposition 4 shows that $G_{ij}(x, t) = \sum_{k=0}^{j-1} W_{i+j-k-1}^*(x, t)W_k(x, t)$. Hence

$$\begin{aligned} \sum_{i=1}^{\infty} G_{ij}(x, t)\mu^{-i} &= \sum_{i=1}^{\infty} W_i^*(x, t)\mu^{-i} \sum_{k=0}^{j-1} W_k(x, t)\mu^{j-k-1} \\ &\quad - \sum_{m=1}^{j-1} \sum_{k=0}^{m-1} W_{m-k}^*(x, t)W_k(x, t)\mu^{j-m-1}, \end{aligned}$$

While from (1.10) we have $\sum_{k=0}^{m-1} W_{m-k}^*(x, t)W_k(x, t) = -W_m(x, t)$ for $m \geq 1$. Thus the first formula of the lemma follows. Similarly we can check the second formula. This proves the lemma.

Let us rewrite (3.17) in terms of $W(x, t, \cdot)$ and $B_{n00}^{(\alpha)}(x, t, \cdot)$. First, applying the lemma to the right-hand side of (3.17), we can rewrite (3.17) into the following equation

$$\begin{aligned} \partial \sum_{i,j=1}^{\infty} G_{ij}(x, t)\mu^{-i}\lambda^{-j}/\partial t_{n00}^{(\alpha)} &= \mu^n E_\alpha \sum_{i,j=1}^{\infty} G_{ij}(x, t)\mu^{-i}\lambda^{-j} - \lambda^n \sum_{i,j=1}^{\infty} G_{ij}(x, t)\mu^{-i}\lambda^{-j}E_\alpha \\ &\quad + W(x, t, \mu)^{-1} \sum_{k=0}^{n-1} (1_r\mu^{n-k-1} + W_1(x, t)\mu^{n-k-2} + \dots + W_{n-k-1}(x, t)E_\alpha \end{aligned}$$

$$\begin{aligned} & \cdot (1_r \lambda^k + W_1^*(x, t) \lambda^{k-1} + \cdots + W_k^*(x, t)) W(x, t, \lambda) \\ & - (\mu^{n-1} + \mu^{n-2} \lambda + \cdots + \mu \lambda^{n-2} + \lambda^{n-1}) E. \end{aligned}$$

Then, multiplying both sides of the last equation by $\lambda - \mu$, and using (1.11) and the formula defining $B_{n00}^{(\alpha)}(x, t, \lambda)$ in (3.15), we finally obtain, after some calculations, the following equation

$$\begin{aligned} & \partial(W(x, t, \mu)^{-1} W(x, t, \lambda)) / \partial t_{n00}^{(\alpha)} \\ & = \mu^n E_x W(x, t, \mu)^{-1} W(x, t, \lambda) - \lambda^n W(x, t, \mu)^{-1} W(x, t, \lambda) E_x \\ & + W(x, t, \mu)^{-1} (B_{n00}^{(\alpha)}(x, t, \lambda) - B_{n00}^{(\alpha)}(x, t, \mu)) W(x, t, \lambda), \end{aligned} \quad (3.18)$$

which is equivalent to (3.17).

We now find that (3.18) is equivalent to the equation for $\rho = \sigma = 0$ in (3.14). To see this, we once more rewrite (3.18) in the following form:

$$\begin{aligned} & (\partial W(x, t, \lambda) / \partial t_{n00}^{(\alpha)} - B_{n00}^{(\alpha)}(x, t, \lambda) W(x, t, \lambda) + \lambda^n W(x, t, \lambda) E_x) W(x, t, \lambda)^{-1} \\ & = (\partial W(x, t, \mu) / \partial t_{n00}^{(\alpha)} - B_{n00}^{(\alpha)}(x, t, \mu) W(x, t, \mu) + \mu^n W(x, t, \mu) E_x) W(x, t, \mu)^{-1}. \end{aligned} \quad (3.19)$$

Then we see immediately that both sides of (3.19) vanish. In fact, (3.19) implies that both its sides are independent of λ and μ , but the definition of $B_{n00}^{(\alpha)}(x, t, \cdot)$ in (3.15) shows that both sides of (3.19) are allowed to contain only the negative powers of λ and μ : Hence they must vanish. Thus we have shown that (3.19) is equivalent to the equation for $\rho = \sigma = 0$ in (3.18), and this proves the equivalence of (3.18) and the latter equation.

Thus we have proved the equivalence of (3.11) and (3.14) in the case where $\rho = \sigma = 0$.

The equivalence in the general case follows from entirely parallel (though more complicated) calculations. The following argument will help us: Using the last formula in (3.12), we can derive from (3.11) the following equation

$$\begin{aligned} & \partial \sum_{i,j=1}^{\infty} G_{ij}(x, t) \mu^{-i} \lambda^{-j} / \partial t_{n\rho\sigma}^{(\alpha)} \\ & = \sum_{\rho'=0}^{\rho} \sum_{\sigma'=0}^{\sigma} \binom{\rho}{\rho'} \binom{\sigma}{\sigma'} y^{\rho-\rho'} \bar{z}^{\rho'} z^{\sigma-\sigma'} (-\bar{y})^{\sigma'} \partial \sum_{i,j=1}^{\infty} G_{ij}(x, t) \mu^{-i} \lambda^{-j} / \partial t_{n+\rho'+\sigma'+00}^{(\alpha)}. \end{aligned}$$

Hence we have

$$\begin{aligned} & \partial(W(x, t, \mu)^{-1} W(x, t, \lambda)) / \partial t_{n\rho\sigma}^{(\alpha)} \\ & = \sum_{\rho'=0}^{\rho} \sum_{\sigma'=0}^{\sigma} \binom{\rho}{\rho'} \binom{\sigma}{\sigma'} y^{\rho-\rho'} \bar{z}^{\rho'} z^{\sigma-\sigma'} (-\bar{y})^{\sigma'} \partial(W(x, t, \mu)^{-1} W(x, t, \lambda)) / \partial t_{n+\rho'+\sigma'+00}^{(\alpha)}. \end{aligned}$$

Now we can apply our previous calculations to the right-hand side of this formula. In this way we can deal with the general case, utilizing the previous discussions.

4. Concluding Remarks

4.1

We have thus far developed our discussions mainly in terms of the matrix ξ . Let us now consider what could be said about $W = \sum_{j=0}^{\infty} W_j \lambda^{-j}$. In fact, in terms of W we can discuss the relation of our method to the method of Mulase [3] and the Riemann–Hilbert problem. We here add some comments on this subejct.

In [3] soliton equations are formulated as the zero-curvature conditions of connections of an infinite dimensional vector bundle whose structure group is a formal Lie group of microdifferential operators, and a kind of “factorization” of formal microdifferential operators is effectively used to construct all the formal power series solutions.

We can develop entirely parallel discussions for the self-dual Yang–Mills equations, replacing microdifferential operators by formal Laurent series of λ . A key lemma in these discussions is that the solution to the intial value problem of (1.4), mentioned in Remark (ii) at the end of Sect. 1, is uniquely characterized by the condition

$$W \cdot \exp(\bar{z}\lambda\partial_y - \bar{y}\lambda\partial_z)W^{(0)-1} \in \mathcal{gl}(r, \mathbb{C}[[x, \lambda]]), \tag{4.1}$$

(i.e. by the condition that $W \cdot \exp(\bar{z}\lambda\partial_y - \bar{y}\lambda\partial_z)W^{(0)-1}$ does not contain the negative powers of λ). We can derive this fact from our previous discussions. In fact, it is easy to show that (4.1) is equivalent to the condition

$$(W_{-j})_{j \in \mathbb{Z}} \tilde{\Xi} = 0, \tag{4.2}$$

where we set $W_j = 0$ for $j < 0$, and (4.2) follows from (2.2), (2.10) and the formula $(W_{-j})_{j \in \mathbb{Z}} \xi = 0$ (this is equivalent to (1.8)) connecting W and ξ . Further, we can recover W directly from (4.2) by

$$(W_{-j})_{j < 0} = -(\tilde{\Xi}_{0j})_{j < 0} \tilde{\Xi}_{(-)}^{-1}, \tilde{\Xi}_{(-)} = (\tilde{\Xi}_{ij})_{i, j < 0}. \tag{4.3}$$

Here the invertibility of $\tilde{\Xi}_{(-)}$ can be easily checked in the same way as in the case of \tilde{H} in Sect. 2. (To tell the truth, all the characteristic matrices of W have such expressions in terms of $\tilde{\Xi}$, since we have $\xi = \tilde{\Xi} \tilde{\Xi}_{(-)}^{-1}$ by virtue of (2.10).) Also we can show that the transformations discussed in Sect. 3 have a similar characterization. Let $W \rightarrow P \circ W$ denote the transformation induced by $\xi \rightarrow \mathbb{P} \circ \xi$, where $P = \sum_{j \in \mathbb{Z}} P_j \lambda^j$. Then $P \circ W$ is uniquely characterized by the condition

$$(P \circ W)PW^{-1} \in \mathcal{gl}(r, \mathbb{C}[[x, \lambda]]). \tag{4.4}$$

We next show that (4.1) is closely related with the Riemann–Hilbert problem. To see this, we first review the setting of the Riemann–Hilbert problem (in a localized situation). Let C be a circle centered at the origin in the Riemann sphere \mathbb{P}^1 , C_+ and C_- respectively the inside and the outside of C , and $u(x, \lambda)$ an invertible $r \times r$ matrix-valued holomorphic function defined in a neighbourhood of $\{0\} \times C$

in $\mathbb{C}^4 \times \mathbb{P}^1$ such that

$$(-\lambda\partial_y + \partial_{\bar{z}})u(x, \lambda) = 0, (\lambda\partial_z + \partial_{\bar{y}})u(x, \lambda) = 0. \quad (4.5)$$

The Riemann–Hilbert problem requires us to find invertible $r \times r$ matrix-valued holomorphic functions $W_+(x, \lambda)$ and $W_-(x, \lambda)$ defined in some neighborhoods of respectively $\{0\} \times (C \cup C_+)$ and $\{0\} \times (C \cup C_-)$ in $\mathbb{C}^4 \times \mathbb{P}^1$ such that

$$W_+(x, \lambda)u(x, \lambda) = W_-(x, \lambda) \quad \text{on } C, W_-(x, \infty) = 1_r. \quad (4.6)$$

If this problem has a solution, then the usual argument (see [6, 11, 12, 16, 17]) shows that $W_+(x, \lambda)$ and $W_-(x, \lambda)$ satisfy (1.1) for some $A_{\bar{y}}$ and $A_{\bar{z}}$, so that we obtain a local holomorphic solution to (0.2). This is the general pattern of applying the Riemann–Hilbert problem to the self-dual Yang–Mills equations. Now, in order to connect this with condition (4.1), let us consider the initial values $W_{\pm}^{(0)}(y, z, \lambda) = W_{\pm}|_{\bar{y}=\bar{z}=0}$ and $u^{(0)}(y, z, \lambda) = u|_{\bar{y}=\bar{z}=0}$. Of course from (4.6) we have $W_+^{(0)}(y, z, \lambda)u^{(0)}(y, z, \lambda) = W_-^{(0)}(y, z, \lambda)$. On the other hand from (4.5) we have $u(x, \lambda) = u^{(0)}(y + \lambda\bar{z}, z - \lambda\bar{y}, \lambda)$ in a sufficiently small neighborhood of $x = 0$. Combining these formulas, we can eliminate $u(x, \lambda)$ from (4.6), and finally obtain the following:

$$W_-(x, \lambda)W_-^{(0)}(y + \lambda\bar{z}, z - \lambda\bar{y}, \lambda)^{-1} = W_+(x, \lambda)W_+^{(0)}(y + \lambda\bar{z}, z - \lambda\bar{y}, \lambda)^{-1}. \quad (4.7)$$

This shows that the Laurent expansion of $W_-(x, \lambda)$ in $C \cup C_-$ (it is a convergent power series of λ^{-1}) satisfies (4.1). In fact, the right-hand side of (4.7) is holomorphic in a neighborhood of $\{0\} \times (C \cup C_+)$, so that it can be expressed as a convergent power series of (x, λ) in the same domain. Thus we find that (4.7) gives a counterpart of condition (4.1) in the framework of the Riemann–Hilbert problem.

Thus we have seen the relation between our method, Mulase’s method and the Riemann–Hilbert problem. Although we can deal with all the formal power series solutions to (0.2) (which constitute a class of solutions much broader than that discussed by the Riemann–Hilbert problem) from our point of view, the relation pointed out above between these methods is still very suggestive.

4.2.

Let us now summarize this paper. First we have introduced an $\infty \times \infty$ matrices $\xi = (\xi_{ij})_{i \in \mathbb{Z}, j < 0}$ consisting of $q\ell(r, \mathbb{C})$ -valued unknown functions ξ_{ij} together with Eqs. (0.3), and established a correspondence of solutions between Eqs. (0.2) and (0.3). By virtue of these discussions we have been able to convert our original problem of describing the structure of solutions to (0.2) into a similar problem for Eqs. (0.3). Next we have shown that both solutions to (0.3) and transformations acting on these solutions can be explicitly constructed by the following simple procedures manipulating $\infty \times \infty$ matrices: $\xi^{(0)} \rightarrow \exp(\bar{z}A\partial_y - \bar{y}A\partial_z)\xi^{(0)} \rightarrow \xi$, $\xi \rightarrow \mathbb{P}\xi \rightarrow \mathbb{P} \circ \xi$. These results, which are very similar to those of Sato [1] for soliton equations, reveals, in terms of ξ , the very simple structure hidden behind the original form of the self-dual Yang–Mills equations. Also, in terms of the unknown function $W = \sum_{j=0}^{\infty} W_j \lambda^{-j}$ of the linear system of Belavin and Zakharov [10],

we have seen that we can develop analogies of the discussions of Mulase [3] devoted to soliton equations, and from this we have found the relation to the Riemann–Hilbert problem.

It is remarkable that the self-dual Yang–Mills equations admit such an approach parallel to Sato’s approach to soliton equations. As for soliton equations, i.e. completely integrable systems in lower dimensions, rich experience has been accumulated for many years. Among them Sato’s theory seems to provide the most universal description of their hidden structure. On the other hand we do not yet know much about higher dimensional completely integrable systems. Hence generalizing Sato’s theory to the higher dimensional cases, for example, is now one of the most challenging problems. Further development of our discussions in this direction may lead to a better understanding of completely integrable systems in higher dimensions.

Appendix

The structure of the matrix $\tilde{\xi}_{(+)}\tilde{\xi}_{(-)}^{-1}$, which appeared in Sect. 2 in the description of the evolution $\xi^{(0)} \rightarrow \xi$, is still less clear, since we there constructed the inverse $\tilde{\xi}_{(-)}^{-1}$ by using the Neumann series (2.5). We here discuss an idea to give more explicit expressions of the components of $\tilde{\xi}_{(+)}\tilde{\xi}_{(-)}^{-1} = (G_{i+1,-j})_{i \geq 0, j < 0}$ by using some determinants of $\infty \times \infty$ matrices.

We begin with a formula of linear algebra. Let $A = (a_{ij})_{1 \leq i, j \leq N}$ be an invertible $N \times N$ matrix, and $f = (f_j)_{1 \leq j \leq N}$ and $g = (g_i)_{1 \leq i \leq N}$ respectively a row vector and a column vector of size N . Then simple calculations of linear algebra yield the formula

$$fA^{-1}g = -\det\left(\begin{array}{c|c} A & g \\ \hline f & 0 \end{array}\right) / \det A. \tag{A.1}$$

We want to establish this formula in the case where $N = \infty$.

To consider this problem, we suppose that A has the form

$$A = 1_N + X, \quad 1_N = (\delta_{ij})_{1 \leq i, j \leq N}, \quad X = (x_{ij})_{1 \leq i, j \leq N}. \tag{A.2}$$

Then we have

$$\begin{aligned} \det(1_N + X) &= 1 + \sum_{k=1}^N \sum_{1 \leq i_1 < \dots < i_k \leq N} \det(x_{i_\alpha i_\beta})_{1 \leq \alpha, \beta \leq k}, \\ \det\left(\begin{array}{c|c} 1_N + X & g \\ \hline f & 0 \end{array}\right) &= -fg + \sum_{k=1}^N \sum_{1 \leq i_1 < \dots < i_k \leq N} \\ &\quad \cdot \det\left(\begin{array}{c|c} (x_{i_\alpha i_\beta})_{1 \leq \alpha, \beta \leq k} & (g_{i_\alpha})_{1 \leq \alpha \leq k} \\ \hline (f_{i_\beta})_{1 \leq \beta \leq k} & 0 \end{array}\right), \end{aligned} \tag{A.3}$$

which are finite dimensional analogues of the well known Fredholm determinant and minors. Hence we may expect that even if $N = \infty$, formula (A.1) holds true by defining the determinants in (A.1) by (A.3), whose right-hand side terms then become infinite series.

In fact, we can justify (A.1) for $\infty \times \infty$ matrices of the form $A = \mathbb{1} + X, f = (f_j)_{1 \leq j < \infty}$ and $g = (g_i)_{1 \leq i < \infty}$ with $\mathbb{1} = (\delta_{ij})_{1 \leq i, j < \infty}$, $X = (x_{ij})_{1 \leq i, j < \infty}$ and $x_{ij}, f_j, g_i \in \mathbb{C}[[x]]$, on the assumption that

$$\text{ord}_x x_{ij} \geq [(i-1)/r] + 1, \text{ord}_x g_i \geq [(i-1)/r] + 1 \text{ for } i, j \geq 1, \quad (\text{A.4})$$

where r is a positive integer and the sign $[\cdot]$ denotes the operation of taking the integral part. Here the inverse of $\mathbb{1} + X$ is defined by the Neumann series $(\mathbb{1} + X)^{-1} = \mathbb{1} + \sum_{k=1}^{\infty} (-X)^k$, whose validity can be easily checked in the same way as the discussions in Sec. 2 by virtue of (A.4). The determinants in (A.1), on the other hand, are defined by directly applying (A.3) to the case $N = \infty$, and the infinite series appearing in the right-hand side of (A.3) actually define certain elements of $\mathbb{C}[[x]]$ by virtue of (A.4). (We can prove the above result by approximately X, f and g respectively by the matrices $X^{(N)}, f^{(N)}$ and $g^{(N)}$ ($N = 1, 2, \dots$) of the form

$$\begin{aligned} X^{(N)} &= (x_{ij}^{(N)})_{1 \leq i, j < \infty}, \quad x_{ij}^{(N)} = x_{ij} (i, j \leq N), = 0 \text{ (otherwise)}, \\ f^{(N)} &= (f_j^{(N)})_{1 \leq j < \infty}, \quad f_j^{(N)} = f_j (j \leq N), = 0 \text{ (otherwise)}, \\ g^{(N)} &= (g_i^{(N)})_{1 \leq i < \infty}, \quad g_i^{(N)} = g_i (i \leq N), = 0 \text{ (otherwise)}. \end{aligned}$$

Note that at the stage of the N^{th} approximations $X^{(N)}, f^{(N)}$ and $g^{(N)}$, the formula which we want to establish is reduced to its counterpart for matrices of finite size, namely (A.1) for $N < \infty$.)

Now let us apply formula (A.1) (which we have just generalized to the case where $N = \infty$), replacing A, f and g respectively by $\tilde{\xi}_{(-)}, ({}^t e_\alpha \tilde{\xi}_{k-1, j})_{j < 0}$ and $(\delta_{i, -m} e_\beta)_{i < 0}$, where $\alpha, \beta = 1, \dots, r, k, m = 1, 2, \dots$, and e_α denotes the column vector $e_\alpha = (\delta_{\alpha\beta})_{1 \leq \beta \leq r}$. Recall here that $\tilde{\xi}_{(-)} = \mathbb{1} + \xi_{(-)}$, and that $\tilde{\xi}_{(-)}$ satisfies (2.4). Hence we can certainly apply formula (A.1) to the case where $A = \tilde{\xi}_{(-)}$. Thus we obtain the following expressions of the components of G_{km} :

$${}^t e_\alpha G_{km} e_\beta = - \det \left(\frac{\tilde{\xi}_{(-)} \quad (\delta_{i, -m} e_\beta)_{i < 0}}{({}^t e_\alpha \tilde{\xi}_{k-1, j})_{j < 0} \quad 0} \right) / \det \tilde{\xi}_{(-)}. \quad (\text{A.5})$$

Furthermore, formulas (A.5) give convergent expressions of the components of G_{km} in the case dealt with by Theorem 7. We can show the convergence of the infinite series defining the determinants in (A.5) as in (A.3), using Hadamard's inequality: Recall that Hadamard's inequality asserts, for any $k \times k$ matrix $(b_{ij})_{1 \leq i, j \leq k}$ with the estimates $|b_{ij}| \leq C c_i d_i$ for $1 \leq i, j \leq k$ for some constants C, c_i and d_i , the following inequality

$$|\det(b_{ij})_{1 \leq i, j \leq k}| \leq (C\sqrt{k})^k \prod_{i=1}^k (c_i d_i). \quad (\text{A.6})$$

Acknowledgement. The author would like to thank Professors M. Sato and H. Komatsu for much help and encouragement, and Doctor M. Mulase for fruitful discussions.

References

1. Sato, M.: Soliton equations as dynamical systems on an infinite dimensional Grassmann manifold. *RIMS Kokyuroku* **439**, 30–46, RIMS, Kyoto Univ. (1981)
2. Date, E., Jimbo, M., Kashiwara, M., Miwa, T.: Transformation groups for soliton equations I–VI, and ditto—Euclidean Lie algebras and reduction of the KP hierarchy —. *Proc. Jpn. Acad.* **57A**, 342–347, 387–392 (1981); *J. Phys. Soc. Jpn.* **50**, 3806–3812, 3813–3818 (1982); *Physica* **4D**, 343–365 (1982); *Publ. RIMS* **18**, 1077–1110, 1111–1120 (1982); Jimbo, M., Miwa, T.: Solitons and infinite dimensional Lie algebras. Preprint, *RIMS-439* (1983)
3. Mulase, M.: Complete integrability of the Kadomtsev–Petviashvili equation. Preprint, MSRI, Berkeley, **053–83** (1983)
4. Segal, G., Wilson, G.: Loop groups and equations of KdV type. Preprint
5. Ueno, K., Takasaki, K.: Toda lattice hierarchy. Preprint, *RIMS-425* (1983); Takasaki, K.: Initial value problem for the Toda lattice hierarchy. Preprint
6. Ward, R.: On self-dual gauge fields. *Phys. Lett.* **61A**, 81–82 (1977)
7. Atiyah, M. F., Ward, R.: Instantons and algebraic geometry. *Commun. Math. Phys.* **55**, 117–124 (1977)
8. Atiyah, M. F., Hitchin, N. J., Drinfeld, V. G., Manin, Yu. I.: Construction of instantons. *Phys. Lett.* **65A**, 185–187 (1978)
9. Atiyah, M. F.: *Geometry of Yang–Mills fields*. Scuola Normale Superiore, Pisa (1979)
10. Belavin, A. A., Zakharov, V. G.: Yang–Mills equations as inverse scattering problem. *Phys. Lett.* **73B**, 53–57 (1978)
11. Corrigan, E. F., Fairlie, D. B., Yates, R. G., Goddard, p.: The construction of self-dual solutions to $SU(2)$ gauge theory. *Commun. Math. Phys.* **58**, 223–240 (1978)
12. Ueno, K., Nakamura, Y.: Transformation theory for anti-self-dual equations and the Riemann–Hilbert problem. *Phys. Lett.* **109B**, 273–278 (1982) Ueno, K., Nakamura, Y.: Transformation theory for anti-self-dual equations. Preprint, *RIMS* **19**, 519–547 (1983)
13. Chau, L. L., Ge, M. L., Wu, Y. S.: Kac–Moody algebra in the self-dual Yang–Mills equations. *Phys. Rev.* **D25**, 1086–1094 (1982) Chau, L. L., Ge, M. L., Sinha, A., Wu, Y. S.: Hidden symmetry algebra for the self-dual Yang–Mills equations. *Phys. Lett.* **121B**, 391–396 (1983)
14. Dolan, L.: A new symmetry group of real self-dual Yang–Mills theory. *Phys. Lett.* **113B**, 387–390 (1982)
15. Pohlmeyer, K.: On the Lagrangian theory of anti-self-dual fields in four dimensional Euclidean space. *Commun. Math. Phys.* **72**, 37–47 (1980)
16. Wu, Y. S.: The group theoretical aspects of infinitesimal Riemann–Hilbert transform and hidden symmetry. *Commun. Math. Phys.* **90**, 461–472 (1983)
17. Chau, L. L., Prasad, M. K., Sinha, A.: Some aspects of the linear systems for self-dual Yang–Mills fields. *Phys. Rev.* **D24**, 1574–1580 (1981)
18. Jimbo, M., Miwa, T.: Monodromy preserving deformation of linear ordinary differential equations with rational coefficients III. *Physica* **4D**, 26–46 (1981)
19. Hauser, I., Ernst, F. J.: Integral equation method for effecting Kinnerseley–Chitre transformations. *Phys. Rev.* **D20**, 362–369 (1979)
20. Ablowitz, M. J., Kaup, D. J., Newell, A. C., Segur, H.: The inverse scattering transform—Fourier analysis for non-linear problem. *Studies in Appl. Math.* **53**, 249–315 (1974)

Communicated by H. Araki

Received October 24, 1983; in revised form December 15, 1983

